

Gain estimation scheme for LPC vocoders with a shape index based on signal envelopes

Patent Number: ☐ US5953697
Publication date: 1999-09-14
Inventor(s): LIN CHIN-TENG (TW); LIN HSIN-AN (TW)
Applicant(s): HOLTEK SEMICONDUCTOR INC (TW)
Requested Patent: ☐ DE19722705
Application Number: US19970851223 19970505
Priority Number(s): TW19960115665 19961219
IPC Classification: G10L9/14
EC Classification: G10L19/08G
Equivalents:

Abstract

A gain estimation method for an LPC vocoder which utilizes shape indexes. The gain is estimated based on the envelope of the speech waveform. The gain is estimated such that the maximum amplitude of the synthetic speech just reaches the speech waveform envelope. The gain during voiced subframes is estimated as the minimum of the absolute value of ratio of the envelope and the impulse response of the LPC filter. The gain during unvoiced subframes is estimated as the minimum of the absolute value of the ratio of the envelope and the noise response of the LPC filter. The method results in a fast technique for estimating the gain.

Data supplied from the esp@cenet database - I2

This Page Blank (uspto)



①9 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑫ **Offenlegungsschrift**
⑩ **DE 197 22 705 A 1**

⑤ Int. Cl.⁶:
G 10 L 7/02
H 03 M 7/30
// H04B 3/23

②1 Aktenzeichen: 197 22 705.8
②2 Anmeldetag: 30. 5. 97
④3 Offenlegungstag: 2. 7. 98

DE 197 22 705 A 1

③0 Unionspriorität:
85115665 19. 12. 96 TW
⑦1 Anmelder:
Holtek Microelectronics Inc., Hsinchu, TW
⑦3 Vertreter:
Haft, von Puttkamer, Berngruber, Czybulka, 81669
München

⑦2 Erfinder:
Lin, Chin-Ten, Tainan, TW; Lin, Hsin-An, I-Lan, TW

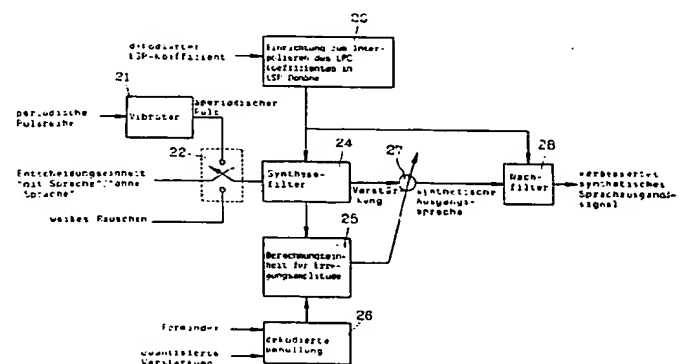
Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

Prüfungsantrag gem. § 44 PatG ist gestellt

⑤4 Verfahren zur Abschätzung der Verstärkung zur Sprachkodierung

⑤7 Die Erfindung betrifft ein Verfahren für ein Schema zur Abschätzung der Verstärkung für die Vokoderkodierung für Sprachvokoderanwendungen mit den folgenden Schritten:

- Erhalten einer dekodierten Hüllkurve, die den Formindex und die quantisierte Verstärkung enthält, durch Anpassen der Eingangssprache aus einem vorbestimmten Code-Lexikon.
- Eingeben entweder eines aperiodischen Pulses oder eines weißen Rauschens direkt in eine Entscheidungseinheit für "mit Sprache"/"ohne Sprache".
- Teilen der Eingangssprache in eine Mehrzahl von Rahmen und Bestimmen, ob jeder Rahmen des Eingangssprachsignals als "mit Sprache oder "ohne Sprache" zu bewerten ist, durch die Entscheidungseinheit.
- Übertragen eines interpolierten linearen Vorhersagekodierungskoeffizienten (LPC) sowohl in das Synthesefilter 24, wie auch in ein Nachfilter 28.
- Übertragen der dekodierten Hüllkurve und des Synthesprachsignals in einer Berechnungseinheit 25 für die Amplitude zur Erzeugung eines Verstärkungssignals.
- Multiplizieren des Verstärkungssignals und des synthetischen Sprachsignals zur Erzeugung eines synthetisierten Sprachausgangssignals.
- Übertragen des synthetisierten Sprachausgangssignals und des interpolierten LPC-Koeffizienten in das Nachfilter 28 zur Erzeugung eines glatten und natürlichen synthetischen Sprachausgangssignals.



DE 197 22 705 A 1

Die Erfindung betrifft ein Verfahren zur Sprachdekodierung mit einem Vokoder und insbesondere ein Verfahren eines Schemas zur Abschätzung der Verstärkung für eine Kodierung eines Vokoders.

Die Technik der linearen voraussagenden Kodierung (LPC) eines Vokoders wird in großem Maße im Zusammenhang mit Synthesatoren-Anwendungen zur Sprachkodierung verwendet. Beispielfhaft wird auf die US-PS 4.910.781 und die US-PS 4.697.261 verwiesen. Die gesamte Offenbarung dieser Patentschriften wird hiernit zum Bestandteil der vorliegenden Patentanmeldung gemacht. Bis heute werden LPC-10 Vokoder in großem Umfang für die Sprachkomprimierung mit einer niedrigen Bitrate verwendet.

Die Fig. 1 zeigt ein Blockschaltbild eines bekannten LPC-Vokoders. Der Vokoder enthält im allgemeinen einen Impulsreihen-Generator 11, einen Zufallsrauschgenerator 12, einen Schalter 13 für "mit Sprache"/"ohne Sprache", eine Verstärkereinheit 14, ein LPC-Filter 15 und eine Einstelleinheit 16 für LPC-Parameter.

Das Eingangssignal des Vokoders wird entweder von dem Impulsreihen-Generator 11 oder dem Zufallsrauschgenerator 12 erzeugt. Der Impulsreihen-Generator 11 kann ein Sprachsignal in der Form einer periodischen Impulsreihe erzeugen, bei dem es sich um das sogenannte Signal "mit Sprache" handelt. Andererseits kann der Zufallsrauschgenerator 12 ein weißes Rauschsignal erzeugen, bei dem es sich um das sogenannte Signal "ohne Sprache" handelt. Entsprechend der richtigen Beurteilung durch den Schalter 13 wird das durch den Impulsreihen-Generator 11 erzeugte Signal mit der periodischen Impulsreihe oder das weiße Rauschsignal, das durch den Rauschgenerator 12 erzeugt wird, zur Verstärkereinheit 14 übertragen. Es wird dann ein LPC-Allpassfilter 15 erregt um ein Ausgangssignal $S(n)$ zu erzeugen, das so skaliert wird, daß es an den Pegel der Eingangssprache angepaßt ist.

Die Sprachentscheidung, die Teilungsperiode, die Filterkoeffizienten und die Verstärkung werden für jeden Sprachrahmen aufdatiert, um Änderungen in der Eingangssprache zu folgen bzw. nachzusteuern. Die Gesamtverstärkung der synthetischen Sprache muß bei praktischen Vokoderanwendungen so eingestellt werden, daß sie an den Pegel der Eingangssprache angepaßt ist. Gegenwärtig gibt es zwei Verfahren zur Bestimmung der Verstärkung, die in großem Umfang angewendet werden. Bei dem ersten Verfahren kann die Verstärkung dadurch bestimmt werden, daß die Energie in dem Sprachsignal an die Energie der linearen vorhergesagten Abtastungen bzw. Proben angepaßt wird. Dies ist in der Tat richtig, wenn richtige Voraussetzungen im Hinblick auf das Erregungssignal für das LPC-System gemacht werden. Einige Voraussetzungen bestehen darin, daß die Vorhersagekoeffizienten α_k in einem tatsächlichen Modell gleich den Vorhersagekoeffizienten α_k in einem realen Modell sind, daß die Energie in dem Erregungssignal $G_u(n)$ für das tatsächliche Modell gleich der Energie in dem Fehlersignal $e(n)$ für das reale Modell ist, daß $u(n) = \delta(n)$ für die gesprochene Sprache ist und daß $u(n)$ für die nicht gesprochene Sprache ein weißes Rauschprozeß mit einem Nullmittelwert und einer Einheitsvarianz ist. Mit dieser Voraussetzung kann die Verstärkung G durch die folgende Gleichung abgeschätzt werden:

$$G = \sqrt{R(0) - \sum_{k=1}^p \alpha_k R(k)} \quad (1)$$

Dabei bezeichnet $R(\cdot)$ die Autokorrelation des Sprachsignales, α_k bezeichnet die LPC-Koeffizienten und p die Vorhersage- bzw. Vorgabeordnung.

Ein weiteres Verfahren zur Verstärkungsberechnung basiert auf dem Effektivwert (RMS) der Abtastungen über dem gesamten Rahmen N der Eingangssprache, der folgendermaßen definiert wird:

$$RMS = \sqrt{\frac{\sum_{n=0}^{N-1} s(n)s(n)}{N}} \quad (2)$$

Für die Rahmen "ohne Sprache" wird die Verstärkung einfach durch RMS abgeschätzt. Für die Rahmen "mit Sprache" wird dieselbe Annäherung auf RMS-Basis verwendet. Die Verstärkung wird jedoch unter Verwendung eines rechteckigen Fensters genauer abgeschätzt, das eine Mehrzahl der gegenwärtigen Teilungsperiode ist. Die nach einem der beiden zuvor genannten Verfahren berechnete Verstärkung wird dann gleichmäßig auf einer logarithmischen Skala unter Verwendung von 7 Bits quantisiert.

Weil der herkömmliche LPC-Vokoder ein System mit einer offenen Schleife ist, reicht ein einfaches Schema zur Abschätzung der Verstärkung nicht aus, um die Amplitude der synthetischen Sprache genau zu bestimmen.

Die Aufgabe der vorliegenden Erfindung besteht darin, ein Verfahren für ein neues Schema zur Abschätzung der Verstärkung für die Vokoderkodierung anzugeben, das glattere und natürlichere Sprachausgangssignale für Vokoderanwendungen erzeugen kann.

Diese Aufgabe wird durch ein Verfahren mit dem Merkmalen des Patentanspruches 1 gelöst.

Vorteilhafterweise können mit dem erfindungsgemäßen Verfahren, das auf der Außenlinie bzw. Kontur der Wellenform der Sprache basiert, die als Umhüllungsform bezeichnet wird, die oben beschriebenen Nachteile vermieden werden.

Die vorliegende Erfindung betrifft ein neues Verfahren für ein Schema zur Abschätzung der Verstärkung für Sprachvokoder, das die folgenden Schritte aufweist:

- Erhalten einer dekodierten Umhüllung, die einen Formindex und eine quantisierte Verstärkung enthält, durch Anpassen einer Eingangssprache aus einem vorbestimmten Code-Lexikon.
- Eingeben entweder eines aperiodischen Pulses oder eines weißen Rauschens direkt in eine Entscheidungseinheit für "mit Sprache"/"ohne Sprache" (voiced/unvoiced unit).
- Unterteilen der Eingangssprache in eine Mehrzahl von Rahmen und Bestimmen, ob jeder Rahmen des Eingangs-

sprachsignales "mit Sprache" oder "ohne Sprache" ist, durch die Entscheidungseinheit.

- d) Übertragen eines interpolierten linearen Vorgabekodierungskoeffizienten (LPC) sowohl in den Synthesefilter, wie auch in einen Nachfilter.
- e) Übertragen der dekodierten Umhüllenden und des Synthese-Sprachsignales in eine Einheit zur Amplitudenberechnung zur Erzeugung eines Verstärkungssignales.
- f) Multiplizieren des Verstärkungssignales und des synthetischen Sprachsignales zur Erzeugung eines synthetisierten Sprachausgangssignales.
- g) Übertragen des synthetisierten Sprachausgangssignales und des interpolierten LPC Koeffizienten in das Nachfilter zur Erzeugung eines glatten und natürlichen verbesserten synthetischen Sprachausgangssignales.

Zum besseren Verständnis werden die Erfindung und deren Ausgestaltungen im Zusammenhang mit den Figuren näher erläutert. Es zeigen:

Fig. 1 ein Blockschaltbild des bekannten Vokoders;

Fig. 2 das Blockschaltbild des erfindungsgemäßen Vokoders und

Fig. 3 die vorbestimmten Form-Codewörter eines 4-Bit-Quantisierers gemäß der vorliegenden Erfindung.

Die vorliegende Erfindung betrifft ein Schema zur Abschätzung der Verstärkung, das auf der Außenlinie der Sprachwellenform basiert, die als Umhüllungsform bezeichnet wird, um die oben beschriebenen Probleme zu lösen.

In der Fig. 2 ist das Blockschaltbild des erfindungsgemäßen Vokoders dargestellt. Der Vokoder umfaßt im allgemeinen einen Vibrator 21, eine Entscheidungseinheit 22 für "mit Sprache"/"ohne Sprache", eine Einrichtung 23 zum Interpolieren des LPC Koeffizienten in Domänen von Linienspektrumpaaren (LSP), ein Synthesefilter 24, das aus einem Allpolfilter und einem Nachentzerrungsfilter besteht, eine Einheit 25 zur Amplitudenberechnung, eine Einrichtung 26 zur Dekodierung der Umhüllung, eine Verstärkungseinheit 27 und ein Nachfilter 28.

Eine durch den Vibrator 21 verlaufende periodische Impulsreihe erzeugt einen aperiodischen Puls an der Entscheidungseinheit 22 "mit Sprache"/"ohne Sprache". Andererseits wird ein weißes Rauschen ebenfalls zu der Entscheidungseinheit 22 gesendet. Gemäß dem erfindungsgemäßen Schema zur Entscheidung zwischen "mit Sprache"/"ohne Sprache" wird ein Rahmen in vier Unterrahmen unterteilt und wird für jeden Unterrahmen auf der Basis einer Anzahl von Parametern, die die normalisierte Korrelation NC, die Energie, den Koeffizienten des Linienspektrumpaars LSP und die Werte des Energieverhältnisses des niedrigen Bandes zum hohen Band (LOH) umfassen, bestimmt, ob der Unterrahmen "mit Sprache" oder "ohne Sprache" ist, um die Genauigkeit des Vokoders in hohem Maße zu vergrößern. Einzelheiten über das Schema zur Entscheidung "mit Sprache"/"ohne Sprache" auf einer 1/4-Basis sind in der deutschen Patentanmeldung 197 21 684.6 desselben Anmelders enthalten.

In ununterbrochenen bzw. andauernden Bereichen der sich langsam ändernden spektralen Charakteristiken kann die rahmenweise Aufdatierung sehr gut funktionieren. In den Übergangsbereichen versagt die rahmenweise Aufdatierung jedoch, wenn Übergänge in den Rahmen fallen. Um sicherzustellen, daß die Ausgangssignale der Übergangsbereiche genauer sind, wird eine bekannte Technik verwendet, um LPC Koeffizienten in der LSP Domäne 23 zu interpolieren, bevor die LPC Koeffizienten zum Synthesefilter 24 gesendet werden. Die Idee besteht darin, eine verbesserte Darstellung des Spektrums dadurch zu erreichen, daß Zwischensätze von Parametern zwischen Rahmen bewertet werden, so daß Übergänge an den Rahmenrändern glatter eingefügt werden, ohne daß die Kodierungskapazität vergrößert wird. Es wurde herausgefunden, daß die Glattheit der verarbeiteten Sprache beträchtlich verbessert wird und daß die Ausgangsqualität der durch schnelle Redner gesprochenen Sprache bemerkenswert verbessert wurde. Um die Berechnungsanzahlen der linearen Interpolation LSP zu verringern, wird der Sprachrahmen in vier Unterrahmen unterteilt. Der in jedem Unterrahmen verwendete LSP Koeffizient wird durch lineare Interpolation der LSP Koeffizienten zwischen dem gegenwärtigen und vorgehenden Rahmen erhalten. Die interpolierten LSP Koeffizienten werden dann in LPC Koeffizienten umgewandelt, die sowohl an das Synthesefilter 24, wie auch an das adaptive Nachfilter 28 gesendet werden.

Die LPC Koeffizienten vom Synthesefilter 24 und die dekodierten Umhüllungssignale, die durch die dekodierte Umhüllung 26 erzeugt werden, werden in die Berechnungseinheit 25 für die Amplitude übertragen, um ein Verstärkungssignal zu erzeugen, das an die Verstärkungseinheit 27 gesendet wird und dann das Nachfilter 28 erregt, um eine verbesserte synthetische Ausgangssprache zu erhalten.

Die Eingänge der dekodierten Hüllkurve 26 weisen die Form einer quantisierten Verstärkung und die normalisierte Form des Index auf. Die Hüllkurvenform und die quantisierten Verstärkungsparameter der synthetischen Sprache werden durch eine Analyse/Synthese-Schleife (analysis-by-synthesis loop) erhalten.

Die Hüllkurvenkodierung wird unter Anwendung einer Annäherung des mittleren quadratischen Fehlers der Verstärkungsform aus einem Code-Lexikon ausgeführt. Durch Minimierung des mittleren quadratischen Fehlers wird die nächstliegende Eingabe aus einem vorbestimmten Code-Lexikon nach der folgenden Gleichung ausgewählt:

$$Error_{env}(i) = \sum_{k=0}^{N-1} (x_i - G_i y_{j,k})^2 \quad (3)$$

$$G_i = \frac{\sum_{k=0}^{N-1} x_k y_{j,k}}{\sum_{k=0}^{N-1} y_{i,k} y_{j,k}} \quad (4)$$

Dabei ist $N=8$ und stellt x_k die Umhüllungsform dar, die zu kodieren ist. $y_{i,k}$ stellt das i^{th} Formcodewort dar. G_i bezeichnet die optimale Verstärkung bei der Anpassung des i^{th} Formcodewortes der Eingangshüllkurve. In der Fig. 3 sind

16 verschiedene Formcodeworte eines 4-Bit Quantisierers gemäß der vorliegenden Erfindung dargestellt. Wenn der optimale Formindex bestimmt wurde, wird die zugeordnete Verstärkung unter Verwendung eines logarithmischen Quantisierers zu 7 Bit quantisiert. Dann werden der Formindex und die quantisierten Verstärkungswerte in die dekodierte Müllkurveneinrichtung 26 gesendet.

Die Verstärkung der Erregung, die so berechnet wird, daß die maximale Amplitude der synthetischen Sprache gerade die dekodierte Hüllkurve erreicht, wird folgendermaßen beschrieben:

a) Unterrahmen "mit Sprache"

Für die Unterrahmen "mit Sprache" weist das Eingangssignal der Entscheidungseinheit 22 für "mit Sprache"/"ohne Sprache" die Form von aperiodischen Impulsen auf. Das Ansprechverhalten des Synthesefilterspeichers (SFMR) wird zuerst von dem vorhergehenden Rahmen herausgefunden. Das Einheits-Impulsansprechen des Synthesefilters 24 an der gegenwärtigen Pulsposition wird dann durch die Berechnungseinheit 25 für die Amplitude berechnet. Die Verstärkung dieses Impulses kann durch folgende Gleichung abgeschätzt werden:

$$\alpha_k = \min_i (abs(Env_{k,i} / imp_res_{k,i})), p_0 \leq i \leq p_0 + r \quad (5)$$

Dabei bezeichnet α_k die K-te Impulsverstärkung $Env_{k,i}$ bezeichnet die dekodierte Hüllkurve für den k-ten Impuls an der Position i. $imp_res_{k,i}$ bezeichnet das Impulsansprechverhalten. p_0 bezeichnet die Impulsposition und r bezeichnet die Suchlänge, die typischerweise 10 beträgt. Wenn die Verstärkung des Pulses herausgefunden ist, wird dieser Puls in das Synthesefilter 24 eingegeben, das ein synthetisches Signal erzeugt. Der SFMR Wert, der gleich dem Produkt des synthetischen Signales und α_k ist, wird in das Nachfilter 28 übertragen, um einen synthetisierten Sprachausgang "mit Sprache" zu erhalten. Das Verfahren wird dann wiederholt, um die Verstärkung des nächsten Impulses herauszufinden.

b) Unterrahmen "ohne Sprache"

Für Unterrahmen "ohne Sprache" weist das Eingangssignal der Entscheidungseinheit 22 die Form eines weißen Rauschens auf. Das Ansprechen des Synthesefilters auf das weiße Rauschen wird zuerst an der Position des gesamten Unterrahmens vollständig berechnet. Dadurch kann die unerwünschte Situation vermieden werden, daß die Amplitude des synthetischen Signales die dekodierte Müllkurve bei diesem Unterrahmen überschreitet. Die Verstärkung des weißen Rauschens an dem gesamten Unterrahmen kann durch die folgenden Gleichung abgeschätzt werden:

$$\beta_j = \min_i (abs(Env_{j,i} / noise_res_{j,i})), w_0 \leq i \leq w_0 + sub_leng \quad (6)$$

Dabei bezeichnet β_j die Verstärkung des weißen Rauschens für den gesamten j-ten Unterrahmen. $Env_{j,i}$ bezeichnet die dekodierte Hüllkurve für dieses weiße Rauschen an der Position i. $noise_res_{j,i}$ bezeichnet das Ansprechen auf das weiße Rauschen. w_0 ist die Anfangsposition jedes Unterrahmens und sub_leng die Unterrahmenlänge. Nachdem die Verstärkung des weißen Rauschens herausgefunden wurde, wird dieses weiße Rauschen dem Synthesefilter 24 zugeführt, das ein synthetisches Signal erzeugt. Der SFMR-Wert, der gleich dem Produkt des synthetischen Signales β_j ist, wird in das Nachfilter 28 übertragen, um ein synthetisiertes Sprachausgangssignal "ohne Sprache" zu erzeugen.

Bei der Ausführung des neuen Schemas zu Abschätzung der Verstärkung für die Vokoderkodierung gemäß der vorliegenden Erfindung werden glattere und natürliche Sprachausgangssignale für Vokoderanwendungen erreicht.

Während die vorliegende Erfindung insbesondere im Zusammenhang mit einem bevorzugten Ausführungsbeispiel erläutert wurde, wird darauf hingewiesen, daß für einen Fachmann im Rahmen der vorliegenden Erfindung zahlreiche Änderungen und Modifikationen möglich sind.

Die Erfindung betrifft ein Verfahren für ein Schema zur Abschätzung der Verstärkung für die Vokoderkodierung für Sprachvokoderanwendungen mit den folgenden Schritten.

- Erhalten einer dekodierten Hüllkurve, die den Formindex und die quantisierte Verstärkung enthält, durch Anpassen der Eingangssprache aus einem vorbestimmten Code-Lexikon.
- Eingeben entweder eines aperiodischen Pulses oder eines weißen Rauschens direkt in eine Entscheidungseinheit für "mit Sprache"/"ohne Sprache".
- Teilen der Eingangssprache in eine Mehrzahl von Rahmen und Bestimmen, ob jeder Rahmen des Eingangssprachsignales als "mit Sprache" oder "ohne Sprache" zu bewerten ist, durch die Entscheidungseinheit.
- Übertragen eines interpolierten linearen Vorhersagekodierungskoeffizienten (LPC) sowohl in das Synthesefilter 24, wie auch in ein Nachfilter 28.
- Übertragen der dekodierten Hüllkurve und des Synthesep Sprachsignales in einer Berechnungseinheit 25 für die Amplitude zur Erzeugung eines Verstärkungssignales.
- Multiplizieren des Verstärkungssignales und des synthetischen Sprachsignales zur Erzeugung eines synthetisierten Sprachausgangssignales.
- Übertragen des synthetisierten Sprachausgangssignales und des interpolierten LPC-Koeffizienten in das Nachfilter 28 zur Erzeugung eines glatten und natürlicheren synthetischen Sprachausgangssignales.

- Leerseite -

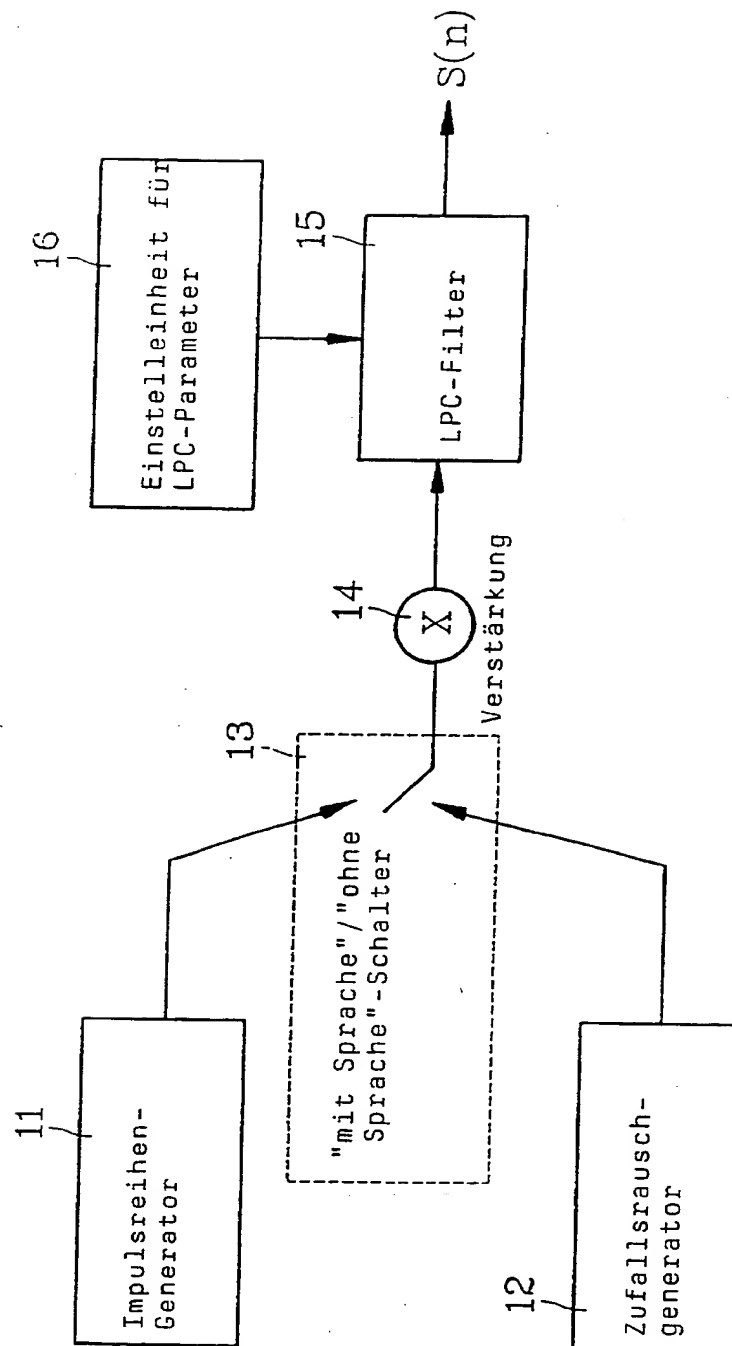


FIG. 1 (STAND DER TECHNIK)

1. Verfahren zur Bestimmung der Verstärkung eines Eingangssprachsignals für Vokoderanwendungen mit den folgenden Schritten:

- a) Erhalten einer dekodierten Hüllkurve durch Anpassung einer Eingangssprache aus einem vorbestimmten Code-Lexikon.
- b) Eingehen eines aperiodischen Pulses durch Eingeben einer periodischen Impulsreihe durch einen Vibrator (21) oder eines weißen Rauschens direkt in eine Entscheidungseinheit (22) zur Entscheidung zwischen "mit Sprache"/"ohne Sprache".
- c) Teilen der Eingangssprache in eine Mehrzahl von Rahmen und Bestimmen, ob jeder Rahmen des Eingangssprachsignals "mit Sprache" oder "ohne Sprache" ist, durch die Entscheidungseinheit (22) und dann Übertragen des Ausgangssignales der Entscheidungseinheit (22) an ein Synthesefilter (24).
- d) Übertragen eines interpolierten linearen Vorhersagekodierungskoeffizienten (LPC) sowohl in das Synthesefilter (24), wie auch in ein Nachfilter (28).
- e) Übertragen der Ausgangssignale der dekodierten Müllkurve (26) und des Synthesefilters (24) in eine Amplitudenberechnungseinheit (25) zur Erzeugung des Verstärkungsausgangssignales.
- f) Multiplizieren des Verstärkungsausgangssignales aus der Amplitudenberechnungseinheit (25) und des synthetischen Ausgangssignales des Synthesefilters (24) durch eine Verstärkungseinheit (27) zur Erzeugung eines synthetisierten Sprachausgangssignales.
- g) Übertragen des synthetisierten Sprachausgangssignales der Verstärkungseinheit (27) und des interpolierten LPC Koeffizienten (23) in das Nachfilter (28) zur Erzeugung eines verbesserten Sprachausgangssignales.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die dekodierte Hüllkurve den Formindex und die quantisierte Verstärkung umfaßt.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß der Formindex und die quantisierte Verstärkung durch eine vorbestimmte Code-Lexikon-Annäherung aus 16 unterschiedlichen Form-Codewörtern mit 4 Bits erhalten werden.

4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß der interpolierte LPC Koeffizient beim Schritt d) durch Interpolieren von LPC Koeffizienten in einer Domäne eines Linienspektrumpaars (LSP) erhalten wird, bevor die LPC Koeffizienten an das Synthesefilter (24) gesendet werden, durch bewerten von Zwischensätzen von Parametern zwischen Rahmen, um die Übergänge an den Rahmenrändern glatter zu machen, ohne die Kodierkapazität zu vergrößern.

5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß das Interpolieren der LPC Koeffizienten in einer Domäne eines Linienspektrumpaars (LSP) durch Unterteilen jedes Sprachrahmens in vier Unterrahmen erhalten wird und daß der verwendete LSP Koeffizient in jedem Sprachrahmen durch lineare Interpolation der LSP Koeffizienten zwischen dem gegenwärtigen Rahmen und den vorhergehenden Rahmen erhalten wird, und daß die interpolierten LSP Koeffizienten dann in die LPC Koeffizienten umgewandelt werden.

6. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß das Verstärkungsausgangssignal der Amplitudenberechnungseinheit (25) beim Schritt e) so berechnet wird, daß die maximale Amplitude der synthetischen Sprache gerade die dekodierte Hüllkurve erreicht und daß die Verstärkung der Unterrahmen "mit Sprache" und "ohne Sprache" getrennt berechnet wird.

7. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß die Verstärkung der Unterrahmen "mit Sprache" durch die folgenden Schritte erhalten wird:

- a) Berechnen einer Einheitspulsansprache des Synthesefilters (24) bei der gegenwärtigen Pulsposition.
- b) Berechnen der Verstärkung des gegenwärtigen Pulses nach der Formel:

$$\alpha_k = \min_i (abs(Env_{k,i} / imp_res_{k,i})), p_0 \leq i \leq p_0 + r$$

wobei α_k die k-te Pulsverstärkung, $Env_{k,i}$ die dekodierte Hüllkurve für den k-ten Puls an der Position i, $imp_res_{k,i}$ das Impulsansprechen, p_0 die Pulsposition und r die Suchlänge (typischerweise 10) bezeichnen.

c) Zuführen des gegenwärtigen Pulses in das Synthesefilter (24) nachdem die Verstärkung des gegenwärtigen Pulses erhalten wurde.

d) Multiplizieren des gegenwärtigen Pulses und α_k zur Erzeugung eines synthetisierten Sprachausgangssignales.

e) Wiederholen der Schritte a) bis d) für den nächsten Puls.

8. Verfahren nach Anspruch 6 oder 7, dadurch gekennzeichnet, daß die Verstärkung der Unterrahmen "ohne Sprache" durch die folgenden Schritte erhalten wird:

- a) Vollständiges Berechnen des Ansprechens auf das weiße Rauschen des Synthesefilters (24) an der Position des gesamten Unterrahmens.
- b) Berechnen der Verstärkung des gesamten Unterrahmens nach der folgenden Gleichung:

$$\beta_j = \min_i (abs(Env_{j,i} / noise_res_{j,i})), w_0 \leq i \leq w_0 + sub_leng$$

wobei α_k die Verstärkung des weißen Rauschens für den gesamten D-ten Unterrahmen, $Env_{j,i}$ die dekodierte Hüllkurve für das weiße Rauschen an der Position i, $noise_res_{j,i}$ das Ansprechen auf das weiße Rauschen, w_0 die Anfangsposition jedes Unterrahmens und sub_leng die Länge des Unterrahmens bezeichnen.

DE 197 22 705 A 1

- c) Zuführen des weißen Rauschens in das Synthesefilter (24) nach dem die Verstärkung des weißen Rauschens erhalten wurde.
d) Multiplizieren des weißen Rauschens und β_j zur Erzeugung eines synthetisierten Sprachausgangssignales.

Hierzu 3 Seite(n) Zeichnungen

5

10

15

20

25

30

35

40

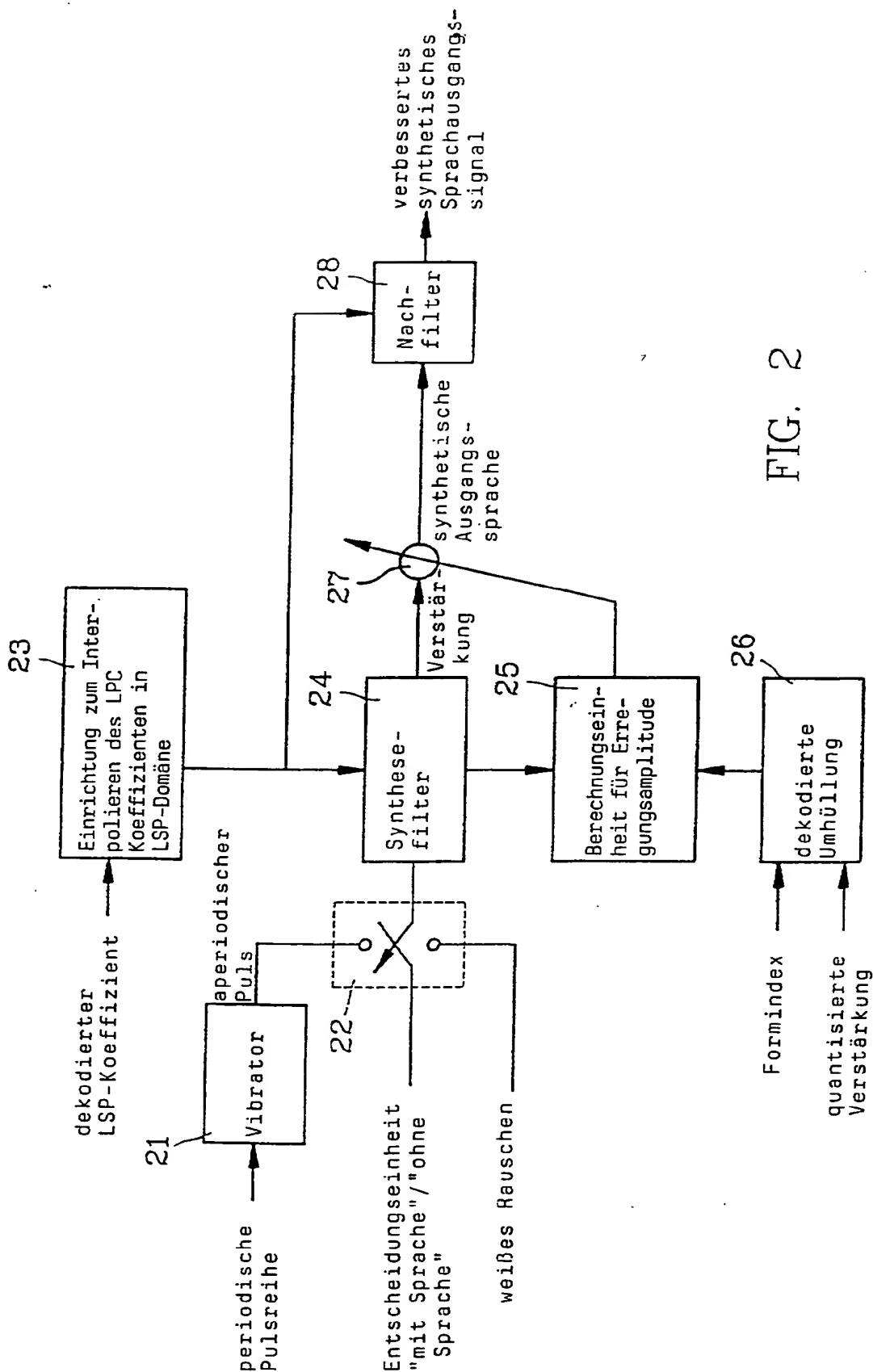
45

50

55

60

65





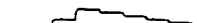
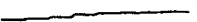











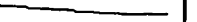
| Index | Hüllkurvencode | | | | | | | | Hüllkurven- form |
|-------|----------------|------|------|------|------|------|------|------|---|
| 0 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |  |
| 1 | 0.60 | 0.66 | 0.73 | 0.79 | 0.86 | 0.93 | 1.00 | 1.00 |  |
| 2 | 0.48 | 0.65 | 0.82 | 1.00 | 0.95 | 0.90 | 0.85 | 0.80 |  |
| 3 | 0.81 | 0.84 | 0.87 | 0.91 | 0.94 | 0.96 | 0.98 | 1.00 |  |
| 4 | 0.51 | 0.58 | 0.64 | 0.72 | 0.78 | 0.86 | 0.93 | 1.00 |  |
| 5 | 1.00 | 0.98 | 0.82 | 0.71 | 0.69 | 0.84 | 0.96 | 0.99 |  |
| 6 | 0.19 | 0.22 | 0.41 | 0.52 | 0.67 | 0.83 | 0.98 | 1.00 |  |
| 7 | 0.40 | 0.52 | 0.64 | 0.76 | 0.88 | 1.00 | 1.00 | 1.00 |  |
| 8 | 0.23 | 0.27 | 0.33 | 0.37 | 0.42 | 0.61 | 0.80 | 1.00 |  |
| 9 | 0.88 | 0.91 | 0.94 | 1.00 | 0.99 | 0.95 | 0.91 | 0.87 |  |
| 10 | 0.04 | 0.06 | 0.21 | 0.36 | 0.52 | 0.68 | 0.84 | 1.00 |  |
| 11 | 1.00 | 0.76 | 0.54 | 0.39 | 0.31 | 0.27 | 0.23 | 0.21 |  |
| 12 | 0.97 | 1.00 | 0.96 | 0.91 | 0.87 | 0.76 | 0.63 | 0.51 |  |
| 13 | 1.00 | 0.95 | 0.91 | 0.87 | 0.83 | 0.79 | 0.75 | 0.75 |  |
| 14 | 1.00 | 0.87 | 0.74 | 0.60 | 0.47 | 0.34 | 0.21 | 0.21 |  |
| 15 | 1.00 | 0.99 | 0.96 | 0.94 | 0.91 | 0.90 | 0.89 | 0.89 |  |

FIG. 3